DØ Note 5883-CONF

Search for the standard model Higgs boson in τ final states

The DØ Collaboration
URL: http://www-d0.fnal.gov

(Dated: March 9, 2009)

We present a search for the standard model Higgs boson using hadronically decaying tau leptons, in 1 fb^{-1} of data collected with the D0 detector at the Fermilab Tevatron $p\bar{p}$ collider. We select two final states: $\tau^{\pm}$ plus missing transverse energy and b jets, and $\tau^{+}\tau^{-}$ plus jets. These final states are sensitive to a combination of associated W/Z boson plus Higgs boson, vector boson fusion and gluon-gluon fusion production processes. The ratio of the combined limit on the Higgs production cross section at the 95% C.L. to the standard model expectation is 27 for a Higgs boson mass of 115 GeV.

Preliminary Results for Winter Conferences 2009

A standard model (SM) Higgs boson with a mass in the range $105-145$ GeV is expected to be produced in 2 TeV $p\bar{p}$ collisions with cross sections of $\mathcal{O}(100$ fb) for associated VH production ($V = W$ or Z) and vector boson fusion (VBF), $q\overline{q} \to VVq'\overline{q}'' \to q'\overline{q}''H$, processes and of $\mathcal{O}(1$ pb) for gluon-gluon fusion (GGF). Previous searches for the SM Higgs boson [1] at the Fermilab Tevatron collider have sought the VH processes with W/Z decays to e/μ and $H \to b\overline{b}$, and the gluon fusion process with $H \to VV^*$ where V decays to e/μ. Thus far, there have been no searches in the case that either the V or H decays to τ leptons. Given the small Higgs boson production cross sections, it is advantageous to use all possible decay modes to increase the search sensitivity. Here, we present a search designed for either of the two final states: $\tau^{\pm}\nu + b\overline{b}$ jets (denoted "$\tau\nu$") or $\tau^+\tau^-+$ jets (denoted "$\tau\tau$"). The analysis is based on 0.94 fb^{-1} ($\tau\nu$) and 1.02 fb^{-1} ($\tau\tau$) of data collected by the D0 experiment [2] at the Fermilab Tevatron collider.

The $\tau\nu$ analysis targets WH production with $W \to \tau\nu$ and ZH production where $Z \to \tau\tau$ but one τ is not identified. The τ is selected through its hadronic decays. We require two jets that are subsequently identified as candidate b quark jets (b tagged). The triggers used for selecting events require jets of high transverse energy, E_T, and large missing transverse energy, $\not\!\!E_T$. The selection of events requires at least one tau candidate decaying to hadrons, at least two jets with transverse momentum $p_T > 15$ GeV, and $\not\!\!E_T$, corrected for the presence of muons and taus, greater than 30 GeV. We reject events containing an electron with $p_T > 15$ GeV or a muon with $p_T > 8$ GeV to maintain statistical independence from other SM Higgs boson searches [1].

The $\tau\tau$ analysis targets VH production with $Z \to \tau^+\tau^-$ and $H \to b\overline{b}$ (denoted "HZ"), $V \to q\overline{q}$ and $H \to \tau^+\tau^-$ ("WH" and "ZH"), VBF with $H \to \tau^+\tau^-$, and GGF with $H \to \tau^+\tau^-$ and at least two associated jets. We identify one of the taus through its decay to $\mu\nu_\tau\overline{\nu}_\mu$ and the other in a hadronic decay mode. The events satisfy a combination of single muon and muon plus jets trigger conditions. They are selected [3] by requiring exactly one muon with $p_T > 12$ GeV, pseudorapidity $|\eta| < 2.0$, and isolated from other tracks and calorimeter activity in a cone surrounding the muon track candidate. We require a hadronic tau candidate and at least two jets. The τ and μ are required to be of opposite charge for the primary event sample. Events containing an electron with $p_T > 12$ GeV are rejected.

We identify three types of hadronic taus, motivated by the decays (1) $\tau^{\pm} \to \pi^{\pm}\nu$, (2) $\tau^{\pm} \to \pi^{\pm}\pi^0\ \nu$, and (3) $\tau^{\pm} \to \pi^{\pm}\pi^{\pm}\pi^{\mp}(\pi^0)\nu$. The identifications [4] are based on the number of associated tracks and activity in the electromagnetic (EM) portion of the calorimeter, both within a cone $\mathcal{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.5$, where ϕ is the azimuthal angle.. The requirements for the $\tau\nu$ ($\tau\tau$) analysis are: for type 1, a single track with $p_T^{\rm trk} > 12$ (15) GeV and no nearby EM energy cluster; for type 2, a single track with $p_T^{\rm trk} > 10$ (15) GeV with an associated EM cluster, and for type 3, at least one track with $p_T^{\rm trk} > 7$ GeV and $\Sigma p_T^{\rm trk} > 20$ GeV and an associated EM cluster. In addition to hadronic τ decays, type 2 taus also contain $\tau \to e$ decays. Type 3 taus are not used in the $\tau\nu$ analysis, and for the $\tau\tau$ channel only those type 3 candidates with unambiguous electric charge sign are retained. The charge is calculated as the sum over the particle charges. A neural network (NN) [4] is formed for each τ type using input variables such as isolation and the transverse and longitudinal shower profiles of the calorimeter energy depositions associated with the tau candidate. Tau preselection is based on the requirement that the output NN value, NN_τ, exceeds 0.3 thus favoring the tau hypothesis. The tau transverse momentum p_T^τ is constructed from the associated track p_Ts, with type-dependent corrections based on the pattern of energy deposit in the associated calorimeter cone. For the three types we require p_T^τ to be greater than 12 (15), 10 (15), or (20) GeV for the $\tau\nu$ ($\tau\tau$) analyses. The $\tau\nu$ analysis subdivides the type 2 taus according to whether the energy deposit is electron-like or hadron-like and the two subsamples are treated separately in assessing the multijet background. For type 2 candidates, we require $0.7 < p_T^{\rm trk}/E_T^\tau < 2$ to remove backgrounds in regions with poor EM calorimetry or due to cosmic rays.

Jets are reconstructed within an $\mathcal{R} = 0.5$ cone in η-ϕ space [5]. Their energies are corrected to the particle level to account for detector effects and missing energy due to semileptonic decays of jet fragmentation products. We require jets to have $p_T > 15$ GeV, $|\eta| < 2.5$, and to be separated by $\mathcal{R} > 0.5$ from τ and μ candidates.

Backgrounds other than those from multijet (MJ) production are simulated using Monte Carlo (MC). We use ALPGEN [6] for $t\overline{t}$ and V+jets production; PYTHIA [7] for WW, WZ and ZZ (diboson) production; and COMPHEP [8] for single top quark production. The ALPGEN events are passed through PYTHIA for parton showering and hadronization. The Higgs boson signal processes are generated using PYTHIA and the CTEQ6L1 [9] leading order parton distribution functions (PDF) for $M_H = 105$ – 145 GeV in 10 GeV steps. We normalize the cross sections to the highest available order calculations for the signal [10] and background [11]. Higgs decays are simulated using HDECAY [12] and tau decays using TAUOLA [13]. All MC events are passed through the standard D0 detector simulation, digitization, and reconstruction programs.

Backgrounds due to MJ production, with fake $\not\!\!E_T$ or misidentified taus are estimated from data samples. For the $\tau\nu$ analysis, an enriched multijet sample is formed by selecting taus with $0.3 < NN_\tau < 0.7$. The contributions from those background processes generated by MC are subtracted to give the BG$_{\tau\nu}$ sample which has negligible Higgs boson signal and provides the shapes of the multijet distributions in the kinematic variables. The normalization is given by the ratio of the number of events in the signal region, $NN_\tau > 0.9$, after subtracting MC backgrounds, to the number of events in the BG$_{\tau\nu}$ sample.

For the MJ background in the $\tau\tau$ analysis, we prepare a multijet background data sample (BG$_{\tau\tau}$) by reversing both

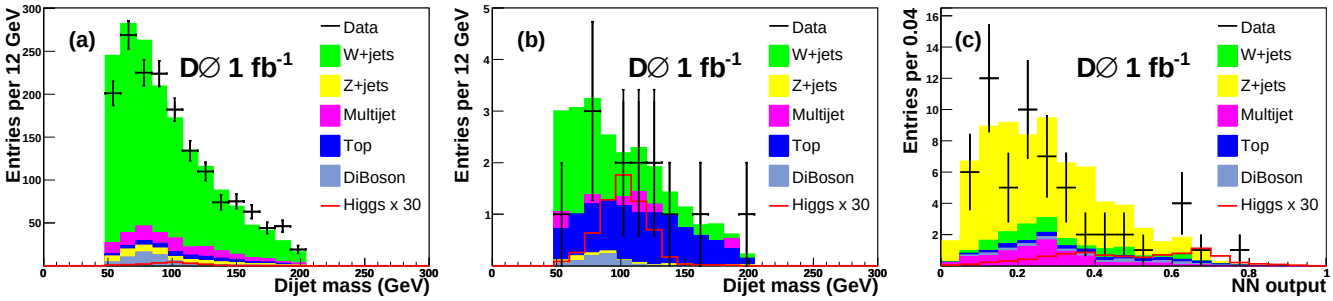

FIG. 1: The dijet mass distribution for all tau types for the $\tau\nu$ analysis (a) before b-tagging, and (b) after the final selection; (c) the combined $NN_{\rm Zjets}$ variable for the low Higgs mass $\tau\tau$ analysis. (color online)

track and calorimeter isolation requirements for the muon and by requiring $NN_\tau < 0.8$. For $\rm BG_{\tau\tau}$, and the sample ($\rm SG_{\tau\tau}$) based on the signal preselection, the MC backgrounds are subtracted, and the same sign (SS) or opposite sign (OS) $\mu-\tau$ charge combinations subsets are formed. The $\rm BG_{\tau\tau}$ sample provides the shape of the multijet background, with the normalization obtained by multiplying the number of SS $\rm SG_{\tau\tau}$ events by the ratio of OS to SS events in the $\rm BG_{\tau\tau}$ sample. These ratios are determined separately for each τ type, and are observed to be close to one and independent of p_T^μ and p_T^τ.

The event sample for the $\tau\nu$ analysis is obtained with additional requirements after the object selections described above: (a) at least two jets with $p_T > 20$ GeV and ≤ 3 jets with $p_T > 15$ GeV; (b) the angle $\Delta\phi(\not\!\!E_T, \not\!\!T_T) < \pi/2$, where $\not\!\!T_T$ is the negative of the transverse component of the net momentum of all tracks in the event; (c) $H_T < 200$ GeV, where H_T is the scalar sum of the p_T of all jets; (d) for hadron-like type 2 taus, the transverse mass, formed from the τ and $\not\!\!E_T$, less than 80 GeV; (e) dijet invariant mass in the range $50 < M_{\rm jj} < 200$ GeV; and (f) the requirement $\Delta\phi(\tau,\not\!\!E_T) < 0.02(\pi-2)(\not\!\!E_T-30)+2$ ($\not\!\!E_T$ in GeV) to reduce contamination due to poorly reconstructed multijet events. To further improve the signal (S) over background (B) separation, we require two jets to be tagged with a NN that discriminates b quark jets and jets from light partons [14]. Figure 1(a,b) shows the $M_{\rm jj}$ distribution before and after b tagging and the event yields are summarized in Table I.

Most of the signal processes sought in the $\tau\tau$ analysis contain light quark jets, so we do not employ b tagging. To further separate signals from backgrounds, we train a dedicated NN for the signal processes (HZ, WH, ZH, VBF) and for each of the main background types ($W+$jets, $Z+$jets, $t\bar{t}$ and MJ). After requiring two jets, the MC GGF samples are small, making NN training unreliable. Since the GGF and VBF processes both involve non-resonant dijet systems, we incorporate the GGF events with the VBF sample when constructing the final limit analysis. The NNs are separately trained for low mass (105, 115 and 125 GeV) and high mass (135, 145 GeV) Higgs bosons, giving 32 NNs in all. Twenty input variables which are well-modeled by the background estimates are considered for each of the NNs. They include transverse or invariant masses of combinations of jets and leptons, $\not\!\!E_T$, angular correlations, and overall event distributions such as H_T and aplanarity[15]. For each signal-background pair, a choice of six or seven variables is made using the criterion that each added variable must give significant improvement in $S/\sqrt{B}$. The same variable choices are made for all Higgs boson masses. All NN input and output variables show good agreement between data and background prediction, and typically provide good discrimination between the signal and background under consideration (though less good discrimination for other backgrounds). We prepare the final event sample by defining the variable $NN_{\rm bg}$ as the maximum of the various signal NN output values for each of the background sources, bg = $t\bar{t}$, $W+$jets, and MJ. We require $NN_{\rm bg} > 0.4$, based on an optimization of the expected Higgs boson cross section limits. After this selection, the NN outputs trained against the $Z+$jets background for all signals are combined by taking their weighted average, $NN_{\rm Zjets}$, over the four signal processes (HZ, WH, ZH, VBF), with weights equal to the relative expected yield for each signal. The $NN_{\rm bg}$ distribution for the final sample is shown in Fig. 1(c), now including the GGF signal events. The signal and background event yields are given in Table I.

Many systematic uncertainties are independent of the final variable distribution shape. Such constant uncertainties for the $\tau\nu$ ($\tau\tau$) analysis are, unless otherwise noted, fully correlated for different backgrounds and analysis channels, and include (a) integrated luminosity, 6.1% (6.1%) [16]; (b) trigger efficiency, 5.5% (3%) (uncorrelated $\tau\nu$ and $\tau\tau$); (c) muon identification, (4.5%); (d) tau identification and track efficiency, 5.1–6.1% (2.5%); (e) tau energy correction, 2.3–2.7% (3.5%); (f) jet identification and reconstruction, 1.7–4.9% (2%); (g) jet energy resolution, (4.5%); (h) jet energy scale (7.5%) [17]; (i) MC background cross sections, 6–18% (6–18%) (these are taken to be uncorrelated among the backgrounds); (j) higher order correction for the $V+$jets cross section, 20% (20%); (k) $V+$ heavy flavor jet cross section correction, 30% (30%); and (l) multijet background, 82–100% (uncorrelated $\tau\nu$ and $\tau\tau$). The possibility that the uncertainties induce a shape dependence on the final limit setting variable is considered. For the $\tau\nu$ analysis, such

TABLE I: Numbers of events at the preselection level and after the final selection (b tagging for $\tau\nu$ and $NN_{\rm bg}$ cut for $\tau\tau$) for all τ types combined, for data, estimated backgrounds and signal at $M_H = 115$ GeV. The V+jets background is given for light parton ("u,d,s,g" = "lp") and heavy flavor ("b,c" = "hf") jets separately. The uncertainties are statistical only.

	$\tau\nu$ analysis		$\tau\tau$ analysis	
Source	Preselection	Final	Preselection	Final
$t\bar{t}$	46.7 ± 0.4	9.5 ± 0.1	30.8 ± 0.3	2.8 ± 0.0
W+ lp	1124 ± 18	0.5 ± 0.0	37.7 ± 2.1	5.1 ± 0.3
W+ hf	308.2 ± 4.8	10.9 ± 0.3	8.2 ± 0.5	0.9 ± 0.1
Z+ lp	49.1 ± 1.5	< 0.2	78.4 ± 0.9	43.8 ± 0.6
Z+ hf	7.8 ± 0.5	0.4 ± 0.0	15.7 ± 1.0	10.1 ± 0.7
Diboson	54.9 ± 1.1	0.7 ± 0.0	6.1 ± 0.5	2.1 ± 0.2
Multijet	122.6 ± 11.2	1.3 ± 0.1	57.2 ± 8.1	6.5 ± 2.8
Sum	1714 ± 22	23.3 ± 0.4	234 ± 9	71.2 ± 3.9
Data	1666	13	220	58
HZ			0.038	0.029
WH	0.543	0.201	0.145	0.106
ZH	0.023	0.015	0.094	0.069
VBF			0.071	0.059
GGF			0.041	0.030
Sum	0.566	0.216	0.389	0.293

TABLE II: Expected and observed 95% C.L. upper limits on the Higgs boson production cross section relative to the SM expected value, for the $\tau\nu$ and $\tau\tau$ analyses separately and combined.

	$\tau\nu$ analysis		$\tau\tau$ analysis		Combined	
M_H (GeV)	exp.	obs.	exp.	obs.	exp.	obs.
105	33	27	38	35	24	20
115	42	35	42	44	28	27
125	62	60	58	61	39	42
135	105	106	87	58	62	48
145	226	211	153	93	118	81

shape dependence is found for the jet energy scale, jet energy resolution, and the b-tagging efficiencies. Alternate shapes are determined by changing the relevant parameter by ± 1 standard deviation from the nominal value and the changed shapes are input to the limit setting program. For the $\tau\tau$ analysis, only the multijet background is found to give an appreciable shape change. It is determined by varying the method for selecting MJ events, reversing either the muon or the tau requirements, but not both, relative to the standard choice.

The upper limits on the Higgs boson cross section are obtained using the modified frequentist method [18]. For the $\tau\nu$ analysis, the test statistic is the negative log likelihood ratio (LLR) derived from the $M_{\rm jj}$ distribution, binned in 10 GeV steps. For the $\tau\tau$ analysis, the LLR is formed from the $NN_{\rm Zjets}$ final neural network variable binned in steps of 0.05. The confidence levels CL_{s+b} (CL_b) give the probability that the signal plus background (background) LLR value from a set of simulated pseudo-experiments is less likely than that observed, at the quoted C.L. The hypothesized signal cross sections are scaled up from their SM values until the value of $CL_s = CL_{s+b}/CL_b$ reaches 0.05 to obtain the limit cross sections at the 95% C.L., both for expected limits (from the pseudo experiments) and observed limits (using data). In the calculation, all contributions to the systematic uncertainty are varied, subject to the constraints given by their estimated values, to give the best fit [19]. The correlations of each systematic uncertainty among signal and/or background processes are accounted for in the minimization.

The expected and observed upper limits are shown in Table II for the two channels separately and combined. At $M_H = 115$ GeV, the observed (expected) 95% C.L. limit is 27 (28) times that predicted in the SM for the seven signal processes considered in the combined $\tau\nu$ and $\tau\tau$ analyses. This is the first limit on SM Higgs production using final states involving hadronically decaying tau leptons. These results contribute to the sensitivity of the the combined Tevataron search for low mass Higgs bosons [20].

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); STFC (United Kingdom); MSMT and GACR (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); CAS and CNSF (China); and the Alexander von Humboldt

Foundation (Germany).

[1] For references to the full set of Higgs searches by the CDF and D0 collaborations, see the TEVNPH working group summary, FERMILAB-PUB-08-270-E.
[2] S. Abachi *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res. A **338**, 185 (1994); V.M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res. A **565**, 463 (2006); V.M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res. A **552**, 372 (2005).
[3] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev Lett. **101**, 241802 (2008).
[4] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev D **71**, 072004 (2005); erratum *ibid* **77**, 039901 (2008).
[5] G.C. Blazey *et al.*, FERMILAB-PUB-00-297 (2000).
[6] M.L. Mangano *et al.*, JHEP **0307**, 001 (2003); we use ALPGEN version 2.05.
[7] T. Sjöstrand *et al.*, arXiv:hep-ph/030815 (2003); we use PYTHIA version 6.319.
[8] E. Boos *et al.*, Phys. Atom. Nucl. **69**, 1317 (2006); E. Boos *et al.*(CompHEP Collaboration), Nucl. Instrum. Methods Phys. Res. A **534**, 250 (2004).
[9] J. Pumplin *et al.*, JHEP **0207**, 012 (2002).
[10] The V+jets cross sections are normalized using MCFM version 3.45 by J. Campbell and R.K. Ellis, Phys. Rev. D **65**, 113007 (2002). The higher order $t\bar{t}$ cross section is taken from N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003).
[11] T. Hahn *et al.*, arXiv:hep-ph/0607308 (2006).
[12] A. Djouadi, J. Kalinowski, and M. Spira, Comp. Phys. Commun. **108**, 56 (1998).
[13] S. Jadach *et al.*, Comp. Phys. Commun. **76**, 361 (1993).
[14] T. Scanlon, FERMILAB-THESIS-2006-43 (2006).
[15] V. Barger, J. Ohnemus, and R.J.N. Phillips, Phys. Rev D **48**, 3953 (1993).
[16] T. Andeen *et al.*, FERMILAB-TM-2365 (2007).
[17] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **101**, 062001 (2008).
[18] A. Read, J. Phys. G: Nucl. Part. Phys. **28**, 2693 (2002); T. Junk, Nucl. Instrum. Methods A **434**, 435 (1999).
[19] W. Fisher, FERMILAB-TM-2386-E (2006).
[20] March 2009 Tevatron CDF-D0 combined Higgs limits.